I read the manuscript titled «A deep dive into genome assemblies of non-vertebrate animals» by Guiglielmoni et al. with great interest. The authors talk about existing methods and algorithms for constructing contiguous and accurate genome assemblies in the context of metazoan genomes. In my opinion, the article is well written and easily understandable by non-specialists. I only have minor concerns that I would like the authors to address if they agree with me.

Introduction

7;894 => 7,894

Sequencing

Figure 1: I understand the intent of this figure, but I find it pretty challenging to read, and points hide other points. One way of fixing this would be to aggregate the data of each category per year and turn it into a boxplot.
«The resulting reads have a length around twenty kilobases (kb)»: In my experience, PacBio reads usually have a mean size around 15kb that can go up to 25kb (see https://www.nature.com/articles/s41597-020-00743-4 as an example).
«The error rate has also been decreasing with the release of new flow cells and the development of more accurate basecallers such as Bonito.» There is also a new protocol called Q20+, which makes it possible to generate reads with a 1% error rate.

Genome assembly

«DBG-based assemblers require highly accurate reads in which errors are only substitutions, with no indels»: why should there be no indels?
«To this end, heterozygous regions are collapsed in order to keep a single sequence for every region in the genome»: this is true if the genome is not very heterozygous. In the other scenario, both haplotypes can often be retrieved, as heterozygous regions are pretty different.

Assembly pre and post-processing

Table 2 - Long reads error correction: NaS is missing.
https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-015-1519-z
Table 2 - Short and long reads polishing: a new tool called HAPO-G has been published recently and is absent from the list. It has been developed explicitly to polish heterozygous genomes but also handles homozygous genomes.
https://academic.oup.com/nargab/article/3/2/lqab034/6262629
Figure 6: Same as Figure 1

Drawbacks of using Hi-C are not presented. As an example, the fact that gap sizes cannot be estimated is not indicated.
«Assembly and pre/post-processing steps are often combined in one tool» makes it look like there is no need to post-process assemblies further, but if the polishing step is only done with long reads, the final quality will not be great.

Phasing assemblies

Hifiasm is another assembler that can phase haplotypes.